

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Analisa Pengelompokan Kelas BPJS Kesehatan Dengan Menggunakan Metode K-Means

**Studi Kasus: Badan Penyelenggara Jaminan Sosial (BPJS) Kota
Pekanbaru**

TUGAS AKHIR

Diajukan Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik Pada
Jurusan Teknik Informatika

Oleh:



ARY BAGUS JIWANDONO
11551102687



**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM**

**RIAU
PEKANBARU**

2021



1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSETUJUAN

**Analisa Pengelompokan Kelas BPJS Kesehatan Dengan
Menggunakan Metode K-Means
Studi Kasus: Badan Penyelenggara Jaminan Sosial (BPJS) Kota
Pekanbaru**


TUGAS AKHIR

Oleh

ARY BAGUS JIWANDONO
11551102687

Telah diperiksa dan disetujui sebagai laporan tugas akhir
di Pekanbaru, pada tanggal 13 Oktober 2021

Pembimbing I,


Dr. Oktalisa, ST, M.Sc
NIP. 19771028 200312 2 004

LEMBAR PENGESAHAN

**Analisa Pengelompokan Kelas BPJS Kesehatan Dengan
Menggunakan Metode K-Means
Studi Kasus: Badan Penyelenggara Jaminan Sosial (BPJS) Kota
Pekanbaru**

TUGAS AKHIR

Oleh

ARY BAGUS JIWANDONO
11551102687

Telah dipertahankan di depan sidang dewan penguji
sebagai salah satu syarat untuk memperoleh gelar sarjana Teknik Informatika
Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau
di Pekanbaru, pada tanggal 13 Oktober 2021

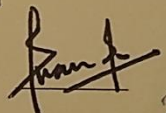
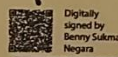
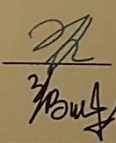


Pekanbaru, 13 Oktober 2021
Mengetahui,
Ketua Jurusan

Iwan Iskandar, M.T.
NIP. 19821216 201503 1 003

DEWAN PENGUJI

Ketua : Iwan Iskandar, M.T.
Sekretaris : Dr. Okfalisa, ST, M.Sc
Penguji I : Benny Sukma Negara, MT
Penguji II : Elvia Budianita, ST, M.Cs



 Digitally signed by
Benny Sukma
Negara


1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta

© H

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Lampiran Surat :

Nomor : Nomor 25/2021

Tanggal : 10 September 2021

SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini:

Nama : ARY BACUS JIWANDONO

NIM : 11551102687

Tempat/Tgl. Lahir : PERAWANG / 6 OKTOBER 1996

Fakultas/Pascasarjana : SAINS dan TEKNOLOGI

Prodi : TEKNIK INFORMATIKA

Judul Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya*:

ANALISA PENGELOMPOKAN KELAS BPJS KESEHATAN DENGAN
MENEGUNAKAN METODE K-MEANS.

Menyatakan dengan sebenar-benarnya bahwa :

1. Penulisan Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya* dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Oleh karena itu Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya* saya ini, saya nyatakan bebas dari plagiat.
4. Apa bila dikemudian hari terbukti terdapat plagiat dalam penulisan Disertasi/Thesis/Skripsi/(Karya Ilmiah lainnya)* saya tersebut, maka saya bersedia menerima sanksi sesuai peraturan perundang-undangan.

Demikianlah Surat Pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 12 NOVEMBER, 2021

Yang membuat pernyataan



ARY BACUS JIWANDONO
NIM : 11551102687

*pilih salah satu sesuai jenis karya tulis



LEMBAR HAK KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dengan ketentuan bahwa hak cipta pada penulis. pada perpustakaan diperbolehkan untuk dicatat, tetapi pengutipan hanya dapat dilakukan dengan seizin penulis dan wajib disertai dengan kebiasaan ilmiah untuk menyebutkan sumber-sumbernya.

Duplikasi atau penerbitan sebagian atau seluruh Tugas Akhir ini harus mendapatkan izin dari Dekan Fakultas Sains Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang ingin meminjamkan Tugas Akhir ini untuk anggotanya diharapkan untuk mengisi data seperti nama, tanda peminjaman dan tanggal penelitian ini dipinjamkan.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bila dalam Tugas Akhir ini tidak mengandung karya yang pernah diajukan sebelumnya untuk memperoleh gelar sarjana di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacukan dalam naskah ini dan disebutkan di daftar pustaka.

Pekanbaru, 2 November 2021

Yang membuat pernyataan,

ARY BAGUS JIWANDONO
11551102687

UIN SUSKA RIAU



LEMBAR PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Syukur, sembah dan sujud hanyalah kepada Allah SWT. Nikmat ilmu dan cintamu telah memberikanku kekuatan dalam menjalani hidup. Memperkenalkan aku dengan ilmu juga memperkenalkanku dengan cinta. Atas karunia serta kemudahan yang telah Engkau berikan akhirnya tugas akhir/skripsi yang sangat sederhana ini dapat terselesaikan dan berjalan dengan baik. Shalawat dan salam selalu terlimpahkan pada baginda Rasulullah Muhammad SAW. Kupersembahkan karya sederhana ini kepada orang yang sangat kusayangi dankukasihi.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Analisa Pengelompokan Kelas BPJS Kesehatan Dengan Menggunakan Metode K-Means

ARY BAGUS JIWANDONO
11551102687

Jurusan Teknik Informatika
Fakultas Sains Teknologi
Universitas Islam Negeri Sultan Syarif Kasim Riau

ABSTRAK

BPJS kesehatan adalah badan usaha yang dibentuk untuk memberikan jaminan kesehatan untuk masyarakat. BPJS Kesehatan sendiri merupakan bentuk transformasi dari layanan yang telah hadir sebelumnya bernama Askes. Sesuai dengan pedoman pelaksanaan JKN yang menyatakan bahwa semua penyakit akan ditanggung kecuali Estetika, infertilitas, alternatif, dan komplementer. BPJS Kesehatan terbagi dalam beberapa golongan yaitu pelayanan kelas 1, kelas 2, dan kelas 3. Pada penelitian ini dilakukan analisis pada penentuan kategori setiap kelas pelayanan tersebut serta membuat pertimbangan golongan kelas menurut pekerjaan, pengasilan, status tempat tinggal, status masyarakat, dan jumlah anggota keluarga. Sering kali dalam penentuan golongan, peserta BPJS Kesehatan tidak memperhitungkan tingkat kemampuan ekonomi dengan golongan yang diambil. Penelitian ini dilakukan agar tidak terjadi kesalahan dalam penentuan golongan yang berakibat terjadinya tunggakan iuran. Penulis akan mencoba menganalisa tempat keempat aspek tersebut dengan menggunakan metode *Clustering K-Means*. Setelah dilakukan perhitungan secara manual, maka dilakukan pengujian untuk memastikan perhitungan analisa tersebut. Langkah pertama yang dilakukan yaitu pengumpulan data, kemudian dilanjutkan dengan pembersihan data. langkah kedua yang dilakukan adalah analisa terhadap proses KDD, analisa atribut, analisa algoritma K-Means, serta pengujian menggunakan *Silhouette Coefficient* dan SSE (*Sum of Square Error*). Hasil yang akan tercipta 3 label yaitu golongan kelas 1, 2, dan 3. Dengan jumlah data yang digunakan adalah 1000 data yang belum memiliki pembagian kelas/golongan. Dalam penelitian ini penulis menggunakan metode clustering K-Means.

Kata Kunci: BPJS, *Sum of Square Error*, *Silhouette Coefficient*, *Clustering*, *K-Means*.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BPJS Health Class Grouping Analysis Using the K-Means Method

ARY BAGUS JIWANDONO

11551102687

Informatics Engineering Departement

Faculty of Science and Technology

Islamic State University Sultan Syarif Kasim Riau

ABSTRACT

BPJS Kesehatan is a business entity formed to provide health insurance for the community. BPJS Health itself is a form of transformation from the service that was present previously called Askes. In accordance with the guidelines for the implementation of JKN which states that all diseases will be covered except for aesthetics, infertility, alternative and complementary. BPJS Kesehatan is divided into several categories, namely Class 1, Class 2, and Class 3. In this study, an analysis was carried out on determining the category of each service class and making consideration of class groups according to occupation, income, status of residence, community status, and number of members. family. Often in determining the group, BPJS Health participants do not take into account the level of economic capacity with the group chosen. This research was carried out so that there would be no mistakes in determining the group which resulted in the occurrence of arrears in contributions. The writer will try to analyze the four four aspects by using the Clustering K-Means method. After the calculation is done manually, then a test is carried out to ensure the calculation of the analysis. The first step taken is data collection, then proceed with data cleaning. The second step taken was an analysis of the KDD process, attribute analysis, K-Means algorithm analysis, and testing using the Silhouette Coefficient and SSE (Sum of Square Error). The results that will be created are 3 groups, namely Class 1, 2, and 3. With the amount of data used, there are 1000 data that do not have class/class division. In this research, the writer uses the C-Means clustering method.

Keywords: BPJS, *Sum of Square Error*, *Silhouette Coefficient*, *Clustering*, *K-Means*.

UIN SUSKA RIAU



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

KATA PENGANTAR



Assalamu 'alaikum wa rahmatullahi wa barakaatuh

Segala puji dan syukur alhamdulillah atas kehadiran Allah Subhanaullahu Wa Ta'ala, yang telah melimpahkan segala nikmat, rahmat, hidayah dan kemudahannya kepada penulis sehingga dapat menyelesaikan tugas akhir yang berjudul “Analisa Pengelompokan Kelas BPJS Kesehatan Dengan Menggunakan Metode K-Means” ini diwaktu yang tepat. Shalawat dan salam kepada Rasulullah Shalallahu „Alaihi Wa Sallam, yang telah membimbing ke jalan yang lurus, sehingga kita dapat merasakan ilmu-ilmu yang memudahkan aktifitas dan ibadah kita sehari-hari.

Tugas akhir ini ini dibuat sebagai salah satu syarat untuk mendapatkan gelar dari Sarjana Teknik jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau. Dalam proses menyelesaikan skripsi ini, penulis telah mendapatkan banyak sekali bimbingan, bantuan, dukungan serta motivasi baik secara langsung maupun tidak langsung. Untuk itu, pada kesempatan ini penulis ingin Ucapan terima kasih kepada:

Bapak Prof. Dr. Khairunnas, M.Ag selaku Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau

Bapak Dr. Hartono, M.Pd selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.

Bapak Iwan Iskandar, M.T sebagai Ketua Jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau.

Ibu Dr. Okfalisa, ST, M.Sc, selaku Pembimbing . Terimakasih untuk kesabaran, waktu, motivasi dan arahan yang telah diberikan sehingga tugas akhir ini dapat terselesaikan.

Bapak Benny Sukma Negara, M.T selaku Penguji I Tugas Akhir yang telah memberikan masukan dan arahan kepada penulis dalam penyusunan tugas akhir ini.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Ibu Elvia Budianita, ST, M.Cs. selaku penguji dua penulis yang telah memberikan kritik serta saran sebagai ilmu yang bermanfaat bagi penulis.

Seluruh Bapak dan Ibu yang mengajar di jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau

Untuk Ayah ku Samiyo dan Almarhumah Ibuku Sumiana terimakasih telah merawatku, memperjuangkan segalanya untuk kebahagiaan ku, ku persembahkan gelar ini untuk kalian ayah dan terkhusus almarhumah ibuku.

Untuk Abangku dan Keluarga besarku, semua yang aku lakukan untuk kalian dan akan aku usahakan.

10. Untuk Keluarga Besar Ikatan Mahasiswa Muhammadiyah UIN Suska terimakasih telah memberikan pengalaman luar biasa untuk ku.

11. Untuk Para Kesatria Birawa Bisma, Very, Pramulia, dan Alfi yang selalu mensupport dan menghibur ku.

12. Teman-temanku Khairul Azmi, Ari Ismanto, Ahmad Muhajir, Yusril Hadi, Ridho Ramidianto, Fikri Syahputra, dan Semua anggota keluarga besar TIF E 2015 dalam berjuang bersama.

13. Semua pihak yang terlibat langsung maupun tidak langsung dalam pelaksanaan tugas akhir ini yang tidak dapat disebutkan satu persatu.

Diharapkan tugas akhir ini bermanfaat bagi penulis khususnya maupun pembaca pada umumnya. Penulis sadar masih banyak sekali kekeliruan dan kekurangan oleh karena itu penulis berharap bisa mendapatkan masukan dari pembaca atas isi tugas akhir ini. Selamat membaca tugas akhir ini saya ucapkan terima kasih.

Pekanbaru, 2 November 2021

Penulis

DAFTAR ISI

BAB I PENDAHULUAN	I-1
1.1 Latar Belakang	I-1
1.2 Rumusan Masalah	I-3
1.3 Batasan Masalah	I-3
1.4 Tujuan Penelitian	I-3
1.5 Sistematika Penulisan	I-3
BAB II LANDASAN TEORI	II-1
2.1 BPJS Kesehatan	II-1
2.2 <i>Data Mining</i>	II-1
2.3 <i>Knowledge Discovery in Database (KDD)</i>	II-2
2.4 <i>Supervised learning dan Unsupervised learning</i>	II-4
2.5 <i>Clustering</i>	II-4
2.6 Algoritma <i>K-means</i>	II-5
2.7 Atribut	II-6
2.8 Metode Elbow (<i>Sum of Square Error</i>)	II-10
2.9 Metode <i>Silhouette Coefficient</i>	II-11
2.10 Penelitian Terkait	II-12
BAB III METODOLOGI PENELITIAN	III-1
3.1 Identifikasi Masalah	III-2
3.2 Studi Literatur	III-2
3.3 Perumusan Masalah	III-2
3.4 Pengumpulan Data	III-2
3.5 Analisa	III-2

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

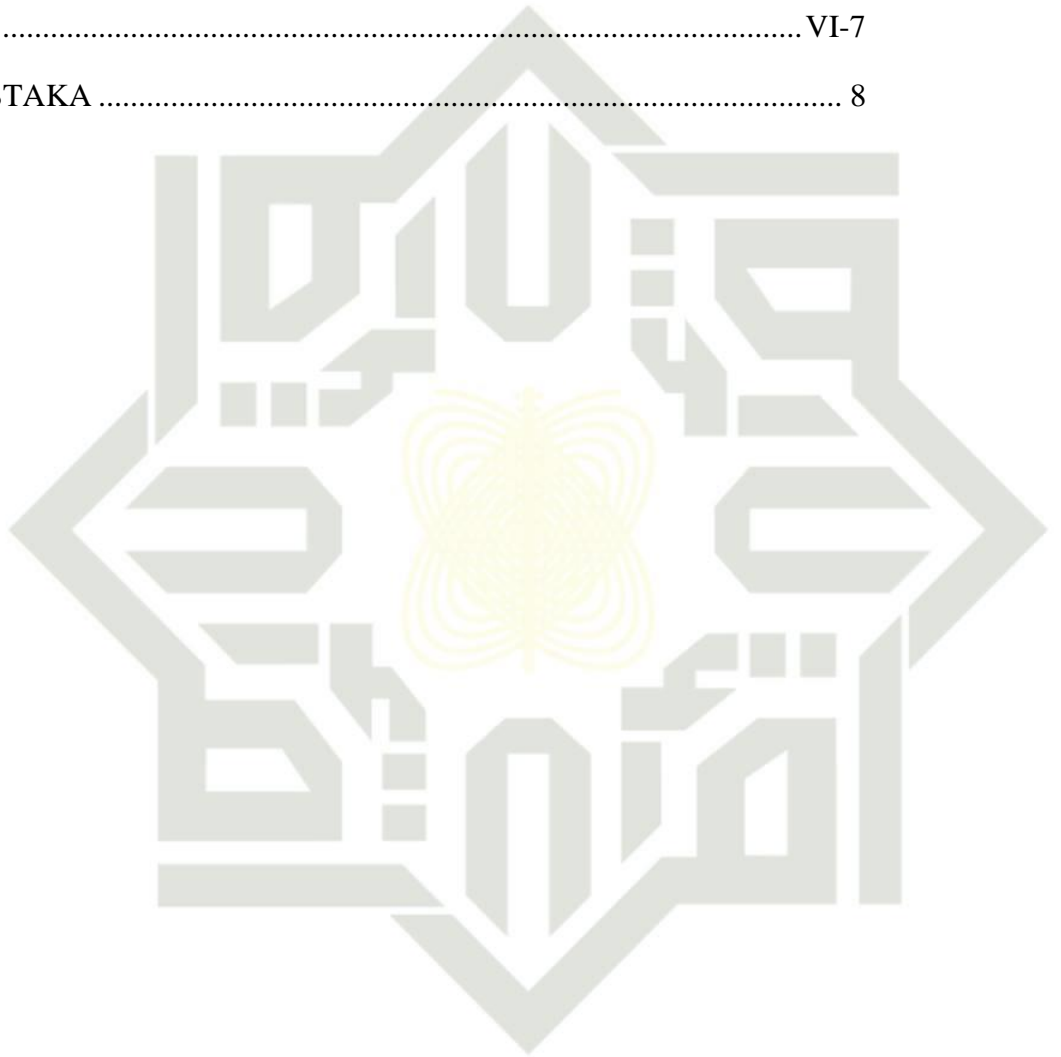
3.5.1	Analisa Proses KDD	III-3
3.5.2	Analisa Atribut	III-4
3.5.3	Analisa Algoritma <i>K-Means</i>	III-4
3.6	Implementasi	III-4
3.7	Pengujian	III-5
3.7.1	Pengujian Elbow SSE (<i>Sum of Square Error</i>)	III-5
3.7.2	Pengujian Silhouette Coefficient	III-5
3.8	Kesimpulan.....	III-5
BAB IV ANALISA DAN PERANCANGAN		IV-1
4.1	Tahapan KDD (<i>Knowledge Discovery in Database</i>)	IV-1
4.1.1	<i>Data Selection</i>	IV-2
4.1.2	<i>Pre-Processing</i>	IV-3
4.1.3	<i>Transformation</i>	IV-3
4.1.4	<i>Data Mining</i>	IV-6
4.2	Penerapan Algoritma <i>K-Means</i>	IV-8
BAB V IMPLEMENTASI DAN PENGUJIAN		V-1
5.1	Implementasi	V-1
5.1.1	Lingkungan Implementasi.....	V- 1
5.1.2	<i>Import Library dan Tools</i>	V-1
5.1.3	Import Data	V-2
5.1.4	Implementasi <i>Preprocessing</i>	V-3
5.1.5	Implementasi <i>Transformation</i>	V-3
5.1.6	Implementasi <i>Clustering</i>	V-4
5.2	Pengujian	V-7
5.2.1	<i>Elbow (Sum of Square of Error)</i>	V-7



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

5.2.2	<i>Silhouette Coefficient</i>	V- 10
5.2.3	Kesimpulan Pengujian	V-11
BAB VI PENUTUP		VI-7
6.1	Kesimpulan.....	VI-7
6.2	Saran	VI-7
DAFTAR PUSTAKA		8



UIN SUSKA RIAU



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR GAMBAR

Gambar 2.1 Data Mining Sebagai Proses dari KDD	3
Gambar 3.1 Flowchart Metodologi Penelitian	1
Gambar 3.2 Flowchart Algoritma K-Means	4
Gambar 4.1 Alur Kerja Algoritma K-Means	7
Gambar 5.1 Import Library	2
Gambar 5.2 Import Data	2
Gambar 5.3 Preprocessing	3
Gambar 5.4 Transformation Tabel	3
Gambar 5.5 Hasil Cluster RStudio	4
Gambar 5.6 Hasil Cluster Dalam Bentuk Tabel	5
Gambar 5.7 Hasil Cluster Dalam Bentuk Grafik	6



DAFTAR TABEL

Tabel 2.1 Spesifikasi Clustering.....	5
Tabel 2.2 Kelas BPJS Kesehatan	7
Tabel 2.3 Tempat Tinggal	7
Tabel 2.4 Jumlah Anggota Keluarga.....	8
Tabel 2.5 Pekerjaan	8
Tabel 2.6 Penghasilan	9
Tabel 2.7 Penelitian Terkait	13
Tabel 4.1 Sampel Data Peserta BPJS Kesehatan	1
Tabel 4.2 Sampel Data Selection	2
Tabel 4.3 Transformasi Nilai Tempat Timggal.....	4
Tabel 4.4 Transformasi Nilai Pekerjaan.....	5
Tabel 4.5 Data Hasil Transformasi	6
Tabel 4.6 Hasil Perhitungan Jarak ke Masing-masing Centroid	10
Tabel 4.7 Hasil Pengelompokan Data Cluster Pada Iterasi Pertama	11
Tabel 4.8 Pembentukan Centroid Baru Iterasi Pertama	14
Tabel 4.9 Hasil Perhitungan Euclidean Distance Pada Iterasi Kedua.....	18
Tabel 4.10 Pembentukan Centroid Baru Iterasi Kedua.....	20
Tabel 4.11 Hasil Perhitungan Euclidean Distance Pada Iterasi Ketiga.....	23
Tabel 4.12 Pembentukan Centroid Baru Iterasi Ketiga.....	25
Tabel 5.1 Detail Hasil Data Uji Coba	7
Tabel 5.2 Selisih Nilai SSE dari Uji Coba 1	8
Tabel 5.3 Selisih Nilai SSE dari Uji Coba 2.....	8
Tabel 5.4 Selisih Nilai SSE dari Uji Coba 3	9
Tabel 5.5 Selisih Nilai SSE dari Uji Coba 4.....	9
Tabel 5.6 Selisih Nilai SSE dari Uji Coba 5.....	9
Tabel 5.7 Hasil Uji Metode Silhouette Coefficient	10

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Asuransi adalah suatu bentuk pertanggung jawaban dalam sebuah perlindungan yang melibatkan dua belah pihak baik itu antara perorangan maupun perusahaan dengan perjanjian dimana pihak pengguna asuransi harus membayar iuran kepada pihak penyedia layanan asuransi dan pihak penyedia layanan asuransi memberikan jaminan sepenuhnya apabila terjadi sesuatu kepada pihak pengguna yang tentunya berdasarkan perjanjian yang telah disepakati.

Pelayanan asuransi saat ini telah memanfaatkan perkembangan teknologi yang sangat pesat ini terlebih lagi penggunaan internet pada saat ini yang merupakan salah satu pendukung pesatnya perkembangan teknologi informasi. Penggunaan media internet memiliki dampak yang sangat menguntungkan bagi kedua belah pihak. Keuntungan dari pihak perusahaan penyedia layanan asuransi yaitu mereka lebih mudah memasarkan layanan serta berbagai macam penawaran dari jasa asuransi mereka. Sedangkan dari pihak pengguna asuransi yaitu pengguna dapat dengan mudah berkomunikasi kepada pihak asuransi apabila terjadi kendala tertentu terlebih lagi untuk masa pandemi ini lebih efektif untuk menggunakan layanan teknologi informasi supaya terhindar dari antrian untuk mengurangi dampak tertularnya penyakit dan juga lebih menghemat waktu yang bisa digunakan kapan saja dan dimana saja.

BPJS (Badan Penyelenggara Jaminan Sosial) merupakan salah satu asuransi milik negara yang memanfaatkan teknologi informasi. Sayangnya belum seluruhnya memanfaatkan teknologi sehingga masih ada kesalahan-kesalahan yang berdampak buruk terhadap perusahaan itu sendiri. Salah satu permasalahannya yaitu mengenai tunggakan iuran peserta BPJS yang disebabkan tidak sesuainya pemilihan kelas dengan tingkat kemampuan ekonomi si peserta. Sehingga banyak peserta memilih kelas berdasarkan keinginannya sendiri yang tidak mempertimbangkan kondisi ekonominya dengan besaran iuran kelas yang



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

diambil yang mengakibatkan tunggakan iuran dimasa yang akan datang. Apabila teknologi dimanfaatkan lagi secara maksimal, bisa saja kesalahan-kesalahan tersebut diminimalisir sehingga banyak kemudahan yang diperoleh perusahaan.

Berdasarkan wawancara penulis dengan kepala bidang keuangan BPJS Kesehatan Kota Pekanbaru bahwa hampir seluruh penduduk Indonesia terdaftar menjadi peserta BPJS Kesehatan yaitu 83% dari total penduduk Indonesia. Dari jumlah tersebut tidak semuanya peserta aktif iuran bulanan BPJS, banyak juga peserta dalam status tidak aktif iuran yang disebabkan tunggakan iuran bulanan. Tunggakan iuran terjadi karena tidak tepatnya sasaran antara tingkat kemampuan peserta dengan kelas yang dipilih. Seharusnya dalam penentuan kelas tersebut kita harus mempertimbangkan beberapa kriteria seperti pengasilan, tanggungan, serta kondisi ekonomi yang lainnya supaya penentuan kelas tepat sasaran dan tidak akan terjadi tunggakan iuran. Dari hasil wawancara dan diskusi dengan kepala bidang keuangan mendapatkan sebuah solusi yaitu mengcluster kenggotaan BPJS Kesehatan dengan mempertimbangkan kriteria-kriteria yang berpengaruh terhadap masalah yang terjadi.

Berdasarkan penelitian sebelumnya yang telah melakukan penelitian terkait dengan algoritma *K-Means* yaitu Siti Syahidatul Helma (2019) yang melakukan penelitian yang berjudul “*Clustering* Pada Data Fasilitas Pelayanan Kesehatan Kota Pekanbaru Menggunakan Algoritma *K-Means*” dengan hasil penelitian yaitu dengan menggunakan algoritma *K-Means* menghasilkan klaster yang optimal dengan menggunakan uji validasi Davies Bouldin Index yang menghasilkan 9 klaster. (Syahidatul Helma et al., 2019)

Kemudian penelitian selanjutnya dilakukan oleh Parasian D P Silitonga dan Romanus Damanik dengan judul penelitian “Implementasi Algoritma *K-Means Clustering* Pada Analisis Penyebaran Penyakit Pasien Pengguna Badan Penyelenggara Jaminan Sosial Kesehatan (BPJS)” dengan hasil penelitian yaitu proses *clustering* yang dilakukan dengan menggunakan metode *K-Means* mendapatkan hasil 4 klaster yang mana 4 klaster tersebut menyimpan data-data

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dari penyakit pasien yang dikelompokkan berdasarkan karakteristik yang sama. (Silitonga & Damanik, 2016)

Dari hasil wawancara dan penjabaran masalah serta penelitian terkait maka penulis akan melakukan penelitian Tugas Akhir yang berjudul “Analisa Pengelompokan Kelas BPJS Kesehatan Dengan Menggunakan Metode K-Means”.

1.2 Rumusan Masalah

Berdasarkan penjabaran latar belakang diatas maka dapat ditarik rumusan masalah yaitu Bagaimana mengcluster data BPJS menggunakan metode *K-means*.

1.3 Batasan Masalah

Batasan masalah dalam penelitian ini adalah:

1. Data yang digunakan berasal dari BPJS Kesehatan Pekanbaru yang merupakan data peserta BPJS se-Pekanbaru.
2. Jumlah *cluster* pada penelitian ini adalah tiga *cluster* yaitu kelas 1, kelas 2, dan kelas 3.
3. Data yang digunakan sebanyak 1000

1.4 Tujuan Penelitian

Tujuan penelitian ini yaitu menganalisa algoritma *K-means* untuk mengklaster kelas peserta BPJS.

1.5 Sistematika Penulisan

Sistematika dalam penulisan penelitian Tugas Akhir ini diurutkan sebagai berikut.

BAB I PENDAHULUAN

Bab ini menjelaskan mengenai pendahuluan dari penulisan proposal dimulai dari latar belakang, rumusan masalah, batasan masalah, tujuan penelitian dan sistematika penulisan.

BAB II LANDASAN TEORI

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Bab ini berisi penjelasan dari teori-teori singkat yang berkaitan dengan topik penelitian serta tentang teori-teori yang mendukung untuk menganalisa algoritma.

BAB III METODOLOGI PENELITIAN

Bab ini berisi rangkaian tahapan-tahapan yang dilalui dalam analisa algoritma dalam penelitian Tugas Akhir ini mulai dari tahapan studi pustaka, perumusan masalah, pengumpulan data, analisa dan perancangan, analisa dan pengujian, hingga kesimpulan dan saran dari hasil penelitian yang diperoleh.

BAB IV ANALISA DAN PERANCANGAN

Bab ini berisi tentang tahapan analisa kebutuhan, kemudian melakukan desain yang dibutuhkan untuk menganalisa dan mengimplementasikan tools yang digunakan.

BAB V IMPLEMENTASI DAN PENGUJIAN

Bab ini berisikan tentang implementasi dari analisa algoritma ke tools yang digunakan, bentuk hasil dari analisa yang sudah dilakukan sebelumnya dan pengujian.

BAB VI PENUTUP

Bab ini menjelaskan mengenai kesimpulan yang dapat diambil dari penelitian ini dan saran yang berguna untuk pengembangan penelitian selanjutnya.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB II LANDASAN TEORI

2.1 BPJS Kesehatan

Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan yang merupakan salah satu lembaga asuransi yang dimiliki oleh pemerintah yang memiliki legalitas yang sah langsung dari Presiden dalam bentuk peraturan Presiden yakni Peraturan Presiden Nomor 12 Tahun 2013. BPJS Kesehatan memiliki tujuan yaitu untuk mensejahterakan seluruh warga negara Indonesia dengan cara memenuhi kebutuhan kesehatan seluruh masyarakat yang tergabung sebagai peserta BPJS Kesehatan baik itu masyarakat yang rutin membayar iuran ataupun masyarakat yang tergabung sebagai anggota BPJS Kesehatan yang iurannya dibayarkan oleh perusahaan swasta ataupun negeri.

Dengan adanya BPJS Kesehatan maka seluruh elemen masyarakat yang ikut menjadi kepesertaan akan mendapatkan layanan kesehatan yang layak dari puskesmas, klinik ataupun rumah sakit yang bekerja sama dengan pemerintah sebagai pemberi layanan BPJS Kesehatan.

2.2 Data Mining

Menurut (Messerli & Devereux, 1983) *data mining* adalah gabungan dari beberapa disiplin ilmu yang menyatukan teknik dari pembelajaran sebuah mesin dengan menggunakan cara pengenalan pola, statistik, *database*, serta visualisasi dalam penanganan masalah dari data yang sangat banyak dan tentunya memiliki *database* yang sangat besar pula.

Mengapa data mining itu diperlukan? Terkadang timbul pertanyaan yang membuat kita ragu. Dari pertanyaan tersebut kita dapat melihat kondisi saat ini yang semakin lama semakin maju, seluruh perusahaan pasti memiliki data-data yang memerlukan pengolahan. Apabila data tersebut diolah secara manual tentu membutuhkan waktu yang sangat lama dan juga membutuhkan tenaga ekstra untuk mengolah data tersebut dan tentunya perlu menggunakan metode-metode

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dalam pengolahan data supaya data tersebut dapat diolah sesuai dengan yang diperlukan. Dengan adanya data mining proses pengolahan data bisa dilakukan dengan cepat dan tepat yang dapat menggunakan metode dan dapat juga memprediksi suatu pola data dalam kasus prediksi.

Data mining sendiri memiliki tujuan yaitu untuk menemukan sebuah pengetahuan yang tersembunyi dari suatu permasalahan dalam *database*. *Data mining* memiliki proses tersendiri dalam menolah data-data yang sangat banyak supaya tersimpan dalam *database*. *Data mining* memiliki pengelompokan dengan masing-masing tugas yang berbeda dalam setiap pengelompokannya. Ada 6 pengelompokan dalam *data mining* (Rika Herman, 2018) yaitu:

- a. Deskripsi, merupakan sebuah pola yang mana dalam proses tersebut kita harus mengguambarkan pola terkait penentuan keselarasan dalam proses *data mining*.
- b. Estimasi, merupakan model dalam *data mining* yang dibangun menggunakan sebuah *record* lengkap yang didalamnya menyediakan nilai berupa variabel sebagai target untuk menentukan nilai prediksi.
- c. Prediksi, merupakan salah satu model pengelompokan yang memiliki tugas untuk menerka sebuah nilai yang disitu belum diketahui kapan nilai itu akan muncul dimasa yang akan datang.
- d. Klasifikasi, dalam klasifikasi terdapat sebuah target yang berbentuk variabel yang telah didefinisikan dalam berbagai macam situasi.
- e. Pengklusteran, memiliki fungsi dalam pengelompokan *record*, pengamatan, atau membentuk keseluruhan objek-objek tersebut menjadi dalam satu kelas..
- f. Asosiasi, merupakan teknik *data mining* yang perannya dapat menemukan hubungan antar kombinasi item dari atribut tertentu.

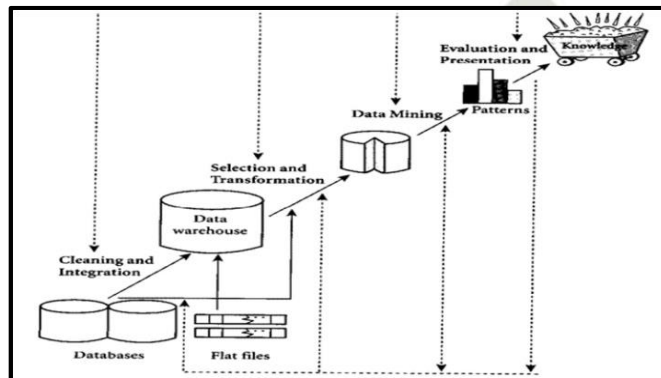
2.3 Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) adalah sebuah proses dalam menemukan informasi yang terdapat dalam *database* (Algoritma et al., 2017). Dengan perkembangan teknologi yang sangat pesat ini kebutuhan-kebutuhan juga

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

semakin banyak pula untuk mengimbangi pesatnya perkembangan teknologi seperti halnya kebutuhan akan tempat penyimpanan data yang sangat besar pula. *Database* atau tempat penyimpanan data haruslah memiliki ruang yang sangat besar yang dianalisa sehingga bermanfaat dan banar-benar menjadi tempat penyimpanan data yang berguna.



Gambar 2.1 Data Mining Sebagai Proses dari KDD

Menurut Aziz Mustafa (Aziz Muustafa et al., 2015) ada beberapa proses secara garis besar dalam KDD yang dijelaskan sebagai berikut:

a. *Data selection*

Data selection merupakan proses pemilihan data dari *database* yang sangat besar dan tentunya memiliki data yang sangat banyak pula yang mana dalam proses tersebut data hasil seleksi disimpan dalam satu berkas yang terpisah.

b. *Pre-processing/ Cleaning*

Setelah dilakukannya proses seleksi pada *database* selanjutnya adalah tahap *cleaning* atau tahap pembersihan yaitu proses dilakukannya pembuangan data-data yang terduplikasi, memperbaiki kesalahan-kesalahan yang terdapat dalam data.

c. *Transformation*

Pada proses ini disebut juga proses *coding* yaitu melakukan transformasi pada data-data yang telah dipilih yang telah sesuai untuk lanjut ke proses *data mining*.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

d. *Data mining*

Pada proses *data mining* inilah dilakukan proses mencari pola terkait informasi-informasi yang menarik dalam data yang sesuai dengan apa yang diperlukan tentunya dengan menggunakan algoritma atau metode tertentu.

e. *Evaluation*

Tahapan terakhir adalah evaluasi. Pada tahap inilah pola-pola yang telah dihasilkan dalam proses *data mining* ditampilkan dalam bentuk yang lebih mudah dimengerti dengan tujuan agar informasi yang telah ditemukan tidak bertentangan dengan hipotesis yang telah dilakukan sebelumnya.

2.4 *Supervised learning dan Unsupervised learning*

Supervised learning merupakan sebuah pendekatan ilmu dalam data mining yang digunakan untuk mengelompokkan data yang mana dalam proses pengelompokan data tersebut memiliki target data yang sudah ada. Sedangkan *Unsupervised learning* merupakan hal sebaliknya dari *supervised learning* yaitu pengelompokan datanya tidak atas dasar data target yang ada atau belum memiliki target. (Sudrajat & Pudjiantoro, 2008)

Dari definisi *supervised learning* dan *unsupervised learning*, algoritma *clustering* merupakan salah satu bentuk dari metode pengelompokan data *unsupervised learning* karena pada saat mengelompokkan data tidak berdasarkan kriteria tertentu atau belum memiliki arahan dan tidak memiliki target.

2.5 *Clustering*

Clustering adalah teknik mengelompokkan data pada *data base* yang mengolah banyak data dalam database tersebut yang berdasarkan kriteria yang telah ditentukan sebelumnya. Pengelompokan data pada metode *clustering* menentukan *cluster* tanpa berdasarkan kelas-kelas tertentu. *Clustering* juga dapat dipakai untuk mengelompokkan data yang kelasnya belum diketahui sama sekali. Bahkan *clustering* dapat digunakan untuk memberikan label pada kelas data yang

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

belum diketahui, sehingga clustering dapat digolongkan kedalam metode *unsupervised learning*. Konsep dari *clustering* sangatlah sederhana yaitu mengelompokkan objek kedalam *cluster* yang memiliki kesamaan, semakin tinggi kesamaan dari objek tersebut maka semakin akurat pula hasil dari *cluster* tersebut.

Dari penjelasan diatas *Clustering* menurut (Han, 2016) lebih spesifik digambarkan sebagai berikut.

Tabel 2.1 Spesifikasi Clustering

Clustering
1. Menganalisis data objek tanpa ada label kelas
2. Label kelas tidak ada atau tidak terlihat pada training data
3. Bertujuan untuk mengelompokkan dan menentukan label kelas dari tiap cluster yang terbentuk
4. Proses clustering berdasarkan pada prinsip: objek yang ada di dalam satu cluster memiliki kemiripan yang tinggi dari pada yang lainnya, tetapi sangat berbeda dengan objek yang ada pada cluster lainnya.

2.6 Algoritma *K-means*

K-means adalah salah satu algoritma dari *clustering*. Algoritma ini termasuk yang cukup sederhana dalam melakukan pengklasteran, yang mana proses pengklasteran algoritma ini dengan cara mempartisi dataset menjadi beberapa klaster. Dalam melakukan klaster algoritma ini mempartisi objek-objek menjadi sebuah kelompok. Pengelompokan tersebut berdasarkan karakteristik yang memiliki kesamaan. Apabila terdapat data yang memiliki karakteristik yang sama maka data tersebut dijadikan satu kelompok klaster dan data yang berbeda dikelompokkan ke dalam kelompok lainnya. (Silitonga & Damanik, 2016)

Dalam melakukan *clustering*, metode *K-means* memiliki beberapa langkah, yaitu sebagai berikut: (Bastian et al., n.d.)

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. Tentukan jumlah *cluster* k
2. Inisialisasikan ke *cluster* k kepusat *cluster*. Penginisialisasian ini dapat dilakukan dengan beberapa cara. Cara yang paling sering digunakan dalam melakukan penginisialisasian ini dengan cara acak (*random*) sehingga pusat-pusat *cluster* diberi nilai awal dengan angka-angka acak
3. Tempatkan seluru data data kedalam clester terdekat. Dua objek dikatakan dekat ditentukan dengan jarak dari kedua objek tersebut. Begitu juga dengan kedekatan data ke cluster.
4. Hitung jarak semua data ke setiap titik pusat cluster yang ada dengan menggunakan rumus *euclidean distance* sebagai berikut:

$$D(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}$$

Keterangan:

$D(i, j)$ = jarak data i ke pusat j

X_{ki} = data ke i pada atribut data ke k

X_{kj} = titik pusat ke j pada atribut ke k

Selanjutnya lakukan penghitungan kembali jarak pusat cluster dengan menggunakan kenggotaan cluster yang sekarang. Pusat cluster merupakan nilai rata-rata dari keseluruhan data yang terdapat pada cluster tertentu. Apabila memungkinkan juga dapat menggunakan median dari cluster yang sama. Oleh karena itu mean atau rata-rata bukanlah satu-satunya ukuran yang bisa digunakan dalam mengklaster data. Setiap objek yang ada ditugaskan lagi untuk menggunakan pusat cluster yang baru. Apabila pada saat proses cluster tidak terjadi perubahan lagi pada pusat cluster, maka proses *clustering* dinyatakan selesai.

2.7 Atribut

Pada proses *clustering* menentukan kelas pada penelitian ini diperlukan atribut yang digunakan sebagai acuan dalam melakukan perhitungan. Atribut yang dipakai adalah sebagai berikut:

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

a. Kelas/ Golongan

Atribut penentuan kelas BPJS Kesehatan menggunakan sifat *cluster* dan guna mengetahui bagaimana pola perilaku perilaku peserta BPJS terhadap kelas yang dipilih maka digunakanlah atribut kelas sebagai inputannya. Berikut atribut kelas BPJS Kesehatan sebagai berikut:

Tabel 2.2 Kelas BPJS Kesehatan

Atribut	Bobot Nilai
Kelas 1	1
Kelas 2	2
Kelas 3	3

b. Tempat Tinggal

Atribut penentuan kelas BPJS Kesehatan dikelompokkan berdasarkan tempat tinggalnya agar lebih mudah untuk menentukan kelas. Tempat tinggal dijadikan sebagai salah satu atribut karena berdasarkan hasil wawancara dengan pihak terkait mengatakan bahwa tempat tinggal adalah syarat terpenting untuk mengetahui apakah peserta memiliki kemungkinan untuk berpindah-pindah tempat tinggal atau tidak supaya lebih mudah untuk mendata alamat si pengguna. Tempat tinggal sendiri memiliki 2 atribut lagi, yaitu Kontrakan dan Rumah sendiri. Detail pemberian bobot dapat dilihat pada tabel berikut:

Tabel 2.3 Tempat Tinggal

Atribut	Bobot Nilai
Kontrak	0
Rumah Sendiri	1

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

c. Jumlah Anggota Keluarga

Atribut penentuan kelas BPJS Kesehatan selanjutnya adalah jumlah anggota keluarga. Atribut ini sangat berpengaruh karena dari kriteria ini dapat dilihat berapa jumlah tanggungan keluarga yang menyesuaikan dengan penghasilan kepala keluarga. Detail pemberian bobot dapat dilihat pada tabel berikut:

Tabel 2.4 Jumlah Anggota Keluarga

Jumlah Anggota Keluarga	Bobot Nilai
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8

d. Jenis Pekerjaan

Atribut jenis pekerjaan ini menentukan akan menentukan jenis pekerjaan apa yang cocok untuk mendapatkan kelas 1,2 atau 3. Karena dari jenis pekerjaan juga dapat menentukan seberapa besar gajinya untuk menanggung anggota keluarganya. Detail pemberian bobot dapat dilihat pada tabel berikut:

Tabel 2.5 Pekerjaan

Pekerjaan	Bobot Nilai
-----------	-------------

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Buruh	0
Guru	1
Pedagang	2
Pegawai	3
Pegawai Swasta	4
Pegawai Negeri	5

e. Penghasilan

Atribut penghasilan adalah yang sangat menentukan dalam penentuan kelas BPJS Kesehatan karena dari penghasilan dapat dianalisa kembali berdasarkan atribut yang sebelumnya seperti jumlah tanggungan keluarga sehingga dapat dilihat seberapa besar pengeluaran dari keluarga tersebut.

Detail pemberian bobot dapat dilihat pada tabel berikut:

Tabel 2.6 Penghasilan

Penghasilan	Bobot Nilai
1.000.000	1.000.000
1.500.000	1.500.000
1.800.000	1.800.000
2.000.000	2.000.000
2.500.000	2.500.000
3.000.000	3.000.000
3.500.000	3.500.000
4.000.000	4.000.000

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4.500.000	4.500.000
5.000.000	5.000.000
5.500.000	5.500.000

2.8 Metode Elbow (*Sum of Square Error*)

Menurut Madhulatha dalam (Muningsih, 2017) menyebutkan bahwa untuk menentukan jumlah cluster yang terbaik yang digunakan sebagai pengujian/evaluasi menggunakan metode *elbow* khususnya SSE (*Sum of Square Error*) merupakan pilihan terbaik karena metode ini melihat hasil persentase dari perbandingan antara jumlah *cluster* yang pada suatu titik akan memebentuk siku.

Untuk mendapatkan perhitungan perbandingan tersebut digunakanlah perhitungan SSE (*Sum of Square Error*) dari tiap-tiap nilai pada *cluster*. Dikarenakan jika nilai dari *cluster* K semakin besar, maka nilai dari SSE akan semakin kecil. Adapun rumus SSE untuk menghitung *cluster* K-Means adalah sebagai berikut:

$$SSE = \sum_{K=1}^K \sum_{x_i \in S_K} \|X_i - C_k\|_2^2$$

Keterangan:

- K : jumlah cluster
- X_i : data ke- i
- C_k : nilai rata-rata dari K .

Algoritma pengujian metode *Elbow* (SSE) untuk penentuan pengujian nilai *cluster* pada metode *K-Means* yaitu sebagai berikut:

1. Inisialisasikan nilai awal K
2. Naikkan angka nilai K
3. Hitunglah hasil SSE dari masing-masing nilai K menggunakan rumus diatas.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4. Amati nilai SSE dari nilai K, lihat manakah nilai K yang terjadi penurunan.
5. Tetapkan bentuk siku dari nilai K

2.9 Metode *Silhouette Coefficient*

Silhouette coefficient merupakan metode yang digunakan untuk memvalidasi sebuah *cluster* dengan menggabungkan nilai kohesi dan separasi. *Silhouette coefficient* dapat digunakan untuk memvalidasi sebuah data, *cluster* tunggal atau bahkan keseluruhan *cluster*. Untuk menghitung nilai *Silhouette Coefficient* terlebih dahulu dilakukan perhitungan *Silhouette Index* dari sebuah data ke- i . perhitungan nilai *Silhouette index* terdapat dua komponen yaitu a_i dan b_i . a_i adalah rata-rata jarak data ke- i dengan semua data lainnya dalam satu *cluster*, sedangkan b_i didapatkan dengan menghitung rata-rata jarak data ke- i terhadap semua data lainnya dalam satu *cluster* yang lain yang tidak dalam satu *cluster* dengan data ke- i , kemudian diambil yang terkecil. (Anggara et al., 2016)

Berikut adalah rumus untuk menghitung a_i

$$a_i^j = \frac{1}{m_j - 1} \sum_{r=1}^{m_j} d(x_i^j, x_r^j)$$

dimana:

- i : index data
- j : Cluster
- $d(x_i^j, x_r^j)$: Jarak data ke- i dengan data ke- r dalam satu *cluster* j
- m_j : Jumlah data dalam *cluster* ke- j
- x : Data

Berikut rumus untuk menghitung b_i^j

$$b_i^j = \min_{n \neq j} \left\{ \frac{1}{m_n} \sum_{r=1}^{m_n} d(x_i^j, x_r^n) \right\}$$

Dimana:

- b : nilai minimum dari rata-rata jarak data ke- i terhadap semua data dari *cluster* yang lain.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

j : Cluster

$d(x_i^j, x_r^j)$: Jarak data ke- i dengan data ke- r dalam satu cluster j

m_n : Jumlah data dalam cluster ke- n

x : Data

untuk mendapatkan *Silhouette Index* (SI) data ke- i menggunakan persamaan sebagai berikut:

$$SI_i^j = \frac{b_i^j - a_i^j}{\max\{a_i^j, b_i^j\}}$$

Dimana:

a : rata-rata jarak data ke- i terhadap semua data lainnya dalam satu cluster.

b : nilai minimum dari rata-rata jarak ke- i terhadap semua data dari cluster yang lain.

$\max\{a_i^j, b_i^j\}$: Nilai maksimum dari nilai a dan b dari satu data

SI : *Silhouette Indeks*

Nilai a_i mengukur seberapa tidak mirip sebuah data dengan cluster yang diikutinya, nilai yang semakin kecil menandakan semakin tepatnya data tersebut berada dalam cluster tersebut. Nilai b_i yang besar menandakan seberapa jauhnya data terhadap cluster yang lain. Nilai SI yang didapat dalam rentang $(-1, 1)$. Nilai SI yang mendekati 1 menandakan bahwa data tersebut semakin tepat berada dalam cluster tersebut. Nilai SI *negative* menandakan bahwa data tersebut tidak tepat berada dalam cluster tersebut (karena lebih dekat ke cluster yang lain).

2.10 Penelitian Terkait

Berikut adalah penelitian yang telah dilakukan sebelumnya terkait dengan penelitian tugas akhir ini.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tabel 2.7 Penelitian Terkait

No.	Judul	Penulis	Tahun	Hasil
1	Pengelompokan Kepemilikan Jaminan Kesehatan Menggunakan Metode Fuzzy C Means Algorithm	Anissa Karmila Islami, Edy Widodo	2017	Hasil proses klaster menggunakan metode fuzzy c-means adalah dari semua kabupaten di Jawa Tengah sebanyak 6 kabupaten yang masuk ke dalam klaster 1, 2 kabupaten dalam klaster 2, 7 kabupaten masuk klaster 3, 12 kabupaten masuk kedalam klaster 4, 1 kabupaten masuk kedalam klaster 5 dan 7 kabupaten masuk kedalam klaster 6.
2	Analisis Pola Penyebaran Penyakit Pasien Pengguna Badan Penyelenggara Jaminan Sosial (Bpjs) Kesehatan Dengan Menggunakan Metode DBSCAN Clustering (Studi Kasus Rumah Sakit Umum Pusat Haji Adam Malik Medan)	Parasian Silitonga	2016	Hasil clustering menggunakan algoritma DBSCAN yang dilakukan terhadap data pasien BPJS kesehatan di Rumah Sakit Adam Malik kota Medan menghasilkan 4 klaster penyimpanan data-data pasien yang mengidap penyakit dengan karakteristik berdekatan.
3	Hierarchical Agglomerative Clustering Untuk Pengelompokan Skripsi Mahasiswa	Herny Februariyanti, Dwi Budi Santoso	2017	Hasil klaser menggunakan metode hierarchial agglomerative menghasilkan semakin rendah nilai threshald maka semakin spesifik pula penentuan judulnya sedangkan semakin tinggi nilai threshald maka semakin luas peula spesifikasi dari judul

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

				skripsi. Threshald yang digunakan 0,25 yang menghasilkan 1 klaster, dan 0,5 menghasilkan 5 klaster.
4	Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang layak Mengikuti Asessment Center Untuk Clustering Program Sdp	In Parlina, Agus Perdana Windarto, Anjar Wanto, M.Ridwan lubis	2018	Hasil cluster dengan menerapkan beberapa kriteria dari daftar Program SDP menggunakan algoritma K-Means dapat diambil pengelompokan dengan rata-rata data program SDP yang dapat melakukan Asessment center lanjutan adalah yang lolos dan hasil klasifikasi program SDP yang Hampir lolos harus memperbaiki Administrasi seperti Kedisiplinan dari bulan juni sampai bulan oktober.
5	Implementasi Algoritma K Means Clustering Pada Analisis Penyebaran Penyakit Pasien Pengguna Badan Penyelenggara Jaminan Sosial Kesehatan (Bpjs)	Parasian D P Silitonga, Romanus Damanik.	2016	Proses klaster menggunakan metode metode K-means mendapatkan kesimpulan yaitu menghasilkan 4 klaster dalam penyimpanan data penyakit pasien yang memiliki karakteristik berdekatan.
6	Analisa Dan Pemanfaatan Algoritma K-Means Clustering Pada Data Nilai Siswa Sebagai Penentuan Penerima Beasiswa	Ari Muzakir	2014	Proses klaster menggunakan metode K-Means memperoleh kesimpulan yaitu proses klaster dengan menggunakan 100 record menggunakan centroid berbeda dapat



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

				menghasilkan klaster yang berbeda juga.
7	Penerapan K-Means Clustering Pada Data Penerimaan Mahasiswa Baru (Studi Kasus : Universitas Potensi Utama)	Fina Nasari, Surya Darma	2015	Proses klaster menggunakan metode K-Means menghasilkan kesimpulan yaitu apabila asal sekolahnya adalah dari SMA maka jurusan yang akan diambil adalah jurusan Sistem informasi sedangkan apabila SMK yang diambil adalah jurusan Teknik Informatika.
8	Clustering pada Data Fasilitas Pelayanan Kesehatan Kota Pekanbaru Menggunakan Algoritma K-Means	Siti Syahidatul Helma, Mustakim, Risma Rustiyan R, Eva Normala	2019	Dari proses klaser menggunakan metode K-means memperoleh kesimpulan yaitu dari hasil uji validasi menggunakan Davies Bouldin Index (DBI) menghasilkan sebanyak 9 klaster. Data yang dominan dalam penentuan klaster di kota Pekanbaru adalah perempuan yang berusia 18 – 20 tahun. Sedangkan di Tampan berdasarkan rata-rata pembelian obat yang besaran harganya yaitu Rp.500.000 dan menggunakan pelayanan BPJS dengan waktu berobat lebih kurang 1 Bulan.
9	Analisa Penentuan Jumlah Cluster Terbaik Pada Metode K-Means Clustering	Ni Putu Eka Merliana, Ernawati, Alb. Joko Santoso	2015	Proses klaster dengan menggunakan metode K-Means memperoleh kesimpulan yaitu dengan metode ini menghasilkan jumlah klaster K sama dengan

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

				jumlah data yang berbeda.
10	Implementasi Metode Klastering K-Means Untuk Mengelompokkan Hasil Evaluasi Mahasiswa	Febrizal Alfarasy Syam	2017	Proses klaster menggunakan metode K-Means memperoleh kesimpulan yaitu pengelompokan mahasiswa berprestasi terdapat di klaster 3 sedangkan yang berpotensi memiliki prestasi terdapat pada klaster 1, yang berpotensi bermasalah terdapat di klaster 2 dan mahasiswa yang bermasalah terdapat di klaster 4.

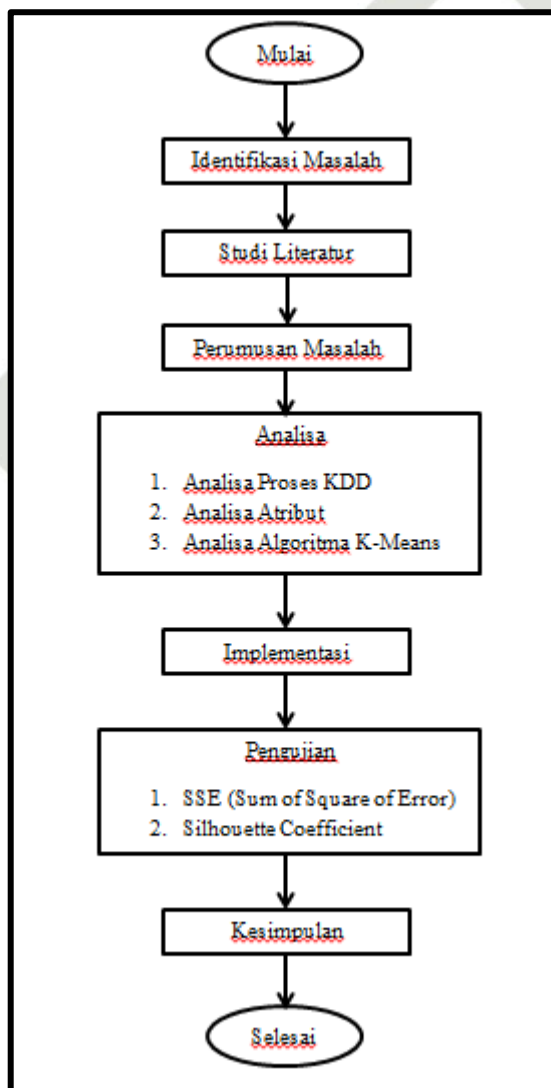
Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB III

METODOLOGI PENELITIAN

Metodologi penelitian merupakan tahapan dari sebuah penelitian yang pengerjaannya dilakukan secara sistematis dan logis yang memiliki tujuan untuk mencapai sebuah target yang telah direncanakan sebelumnya. Berikut ini merupakan flowchart dari penelitian ini.



Gambar 3.1 *Flowchart* Metodologi Penelitian



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

© Hak Cipta dilindungi Undang-Undang
UIN Suska Riau
State Islamic University of Sultan Saifuddin Kasim Riau

3.1 Identifikasi Masalah

Mengidentifikasi suatu permasalahan merupakan tahapan paling awal dalam melakukan sebuah penelitian. Tahapan awal dalam mengidentifikasi permasalahan yaitu dengan melakukan kegiatan untuk mendapatkan gambaran dari permasalahan dan mencari solusi dari permasalahan tersebut. Pada penelitian kali ini permasalahan yang diangkat adalah Bagaimana mengimplementasikan algoritma K-Means dalam mengklaster data BPJS.

3.2 Studi Literatur

Studi literatur merupakan proses mencari referensi teori yang berhubungan dengan penelitian yang dilakukan baik itu dari buku ataupun jurnal-jurnal terkait. Studi literatur ini bertujuan untuk mengetahui teori, metode serta konsep-konsep yang akan digunakan dalam penelitian.

3.3 Perumusan Masalah

Perumusan masalah merupakan tahapan yang dilakukan setelah tahapan studi literatur. Berdasarkan hasil yang diperoleh dari studi literatur yang dilakukan dalam penelitian ini, maka dapat disimpulkan permasalahan yang akan dijadikan sebagai penelitian Tugas akhir ini yaitu menganalisa algoritma K-Means untuk clustering penentuan kelas BPJS Kesehatan.

3.4 Pengumpulan Data

Pengumpulan data merupakan hal yang sangat penting yang dilakukan dalam penelitian karena data merupakan sesuatu yang paling utama untuk mendukung penelitian Tugas Akhir ini. Data yang didapatkan adalah data dalam bentuk file berekstensi .x/s yang diperoleh langsung dari kantor BPJS Kesehatan Cabang Kota Pekanbaru sebanyak 1000 data.

3.5 Analisa

Analisa merupakan tahapan yang dilakukan setelah tahapan pengumpulan data. Pada tahapan analisa ini peneliti melakukan analisa terhadap proses KDD, analisa terhadap atribut-atribut apa saja yang digunakan, lalu analisa terhadap algoritma yang digunakan dalam penelitian ini.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3.5.1 Analisa Proses KDD

Pada tahapan ini dijelaskan bagaimana proses analisa serta tahapan-tahapan dalam melakukan klaster kelas BPJS Kesehatan menggunakan algoritma *K-Means*. Adapun tahapan proses KDD adalah sebagai berikut:

1. *Data Selection*

Tahapan ini dilakukan untuk menentukan atribut apa saja yang digunakan dalam penelitian yang terdapat pada data-data yang tersedia. Atribut-atribut yang dipilih harus memiliki keterkaitan antar atribut agar lebih memudahkan dalam melakukan pengklasteran.

2. *Cleaning/ Pre-Processing*

Tahapan ini dilakukan untuk melakukan pembersihan data terlebih dahulu apakah masih terdapat data-data yang kosong, duplikat, ataupun data yang tidak konsisten.

3. *Transformation*

Data yang telah melalui proses *cleaning* selanjutnya akan ditransformasikan dalam bentuk data yang disesuaikan dengan penerapan pada algoritma yang akan digunakan nanti.

4. *Data Mining*

Proses dalam melakukan pencarian suatu informasi yang bermanfaat yang sesuai dengan tujuan pada kumpulan data yang sangat banyak menggunakan algoritma *K-Means*. Hasil yang diperoleh dari penerapan algoritma tersebut adalah *cluster* data kelas kepesertaan BPJS Kesehatan.

5. *Evaluation*

Tahapan terakhir ialah mengevaluasi hasil dari proses data mining dalam pengelompokan kelas BPJS. Pada tahap ini diuji apakah proses pengklasteran sudah tepat sesuai dengan atribut-atribut yang digunakan dalam proses pengklasteran yang tentunya menggunakan metode dan hipotesis yang telah dilakukan sebelumnya.

Hak Cipta Dilindungi Undang-Undang

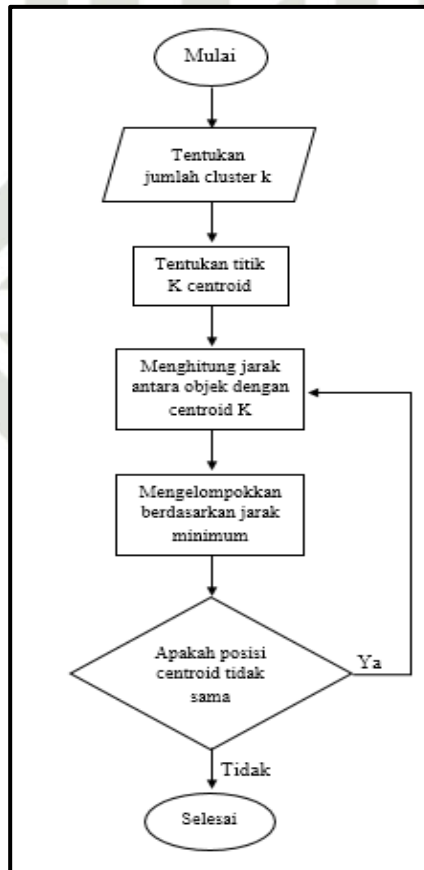
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3.5.2 Analisa Atribut

Analisa atribut merupakan sebuah objek yang dianalisa yang mana menjadi acuan perhitungan penentuan kelas dalam penelitian ini. Pada penelitian ini menggunakan beberapa atribut, yaitu kelas BPJS, tempat tinggal, jumlah anggota keluarga, pekerjaan, dan penghasilan perbulan.

3.5.3 Analisa Algoritma *K-Means*

Secara umum penerapan algoritma *K-Means* untuk metode *clustering* dapat digambarkan dalam bentuk *flowchart* berikut ini.



Gambar 3.2 *Flowchart* Algoritma *K-Means*

3.6 Implementasi

Implementasi atau penerapan algoritma ini terdiri dari lingkungan perangkat keras (*hardware*) dan perangkat lunak (*software*). Berikut adalah penjelasan dari lingkungan implementasi tersebut.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

© Hak cipta milik UIN Suska Riau

2

1. Perangkat Keras (*Hardware*)

- a. *Processor* : Intel Core i3
- b. *RAM* : 4 GB
- c. *Harddisk* : 500 GB

2. Perangkat lunak (*Software*)

3. *Operating System* : Microsoft Windows 7 64 bit
4. *Library* : *library(cluster)*, *library(factoextra)*,
library(tidyverse)
5. *Tools* : RStudio

3.7 Pengujian

Tahap pengujian ini merupakan suatu tahapan guna menguji algoritma yang telah selesai dibuat pada implementasi sebelumnya dan pengujian hasil.

3.7.1 Pengujian Elbow SSE (*Sum of Square Error*)

Tahap pengujian ini dilakukan untuk memperoleh kluster terbaik dalam analisa proses kluster dengan melihat hasil persentase perbandingan dari hasil cluster yang telah ditemukan.

3.7.2 Pengujian Silhouette Coefficient

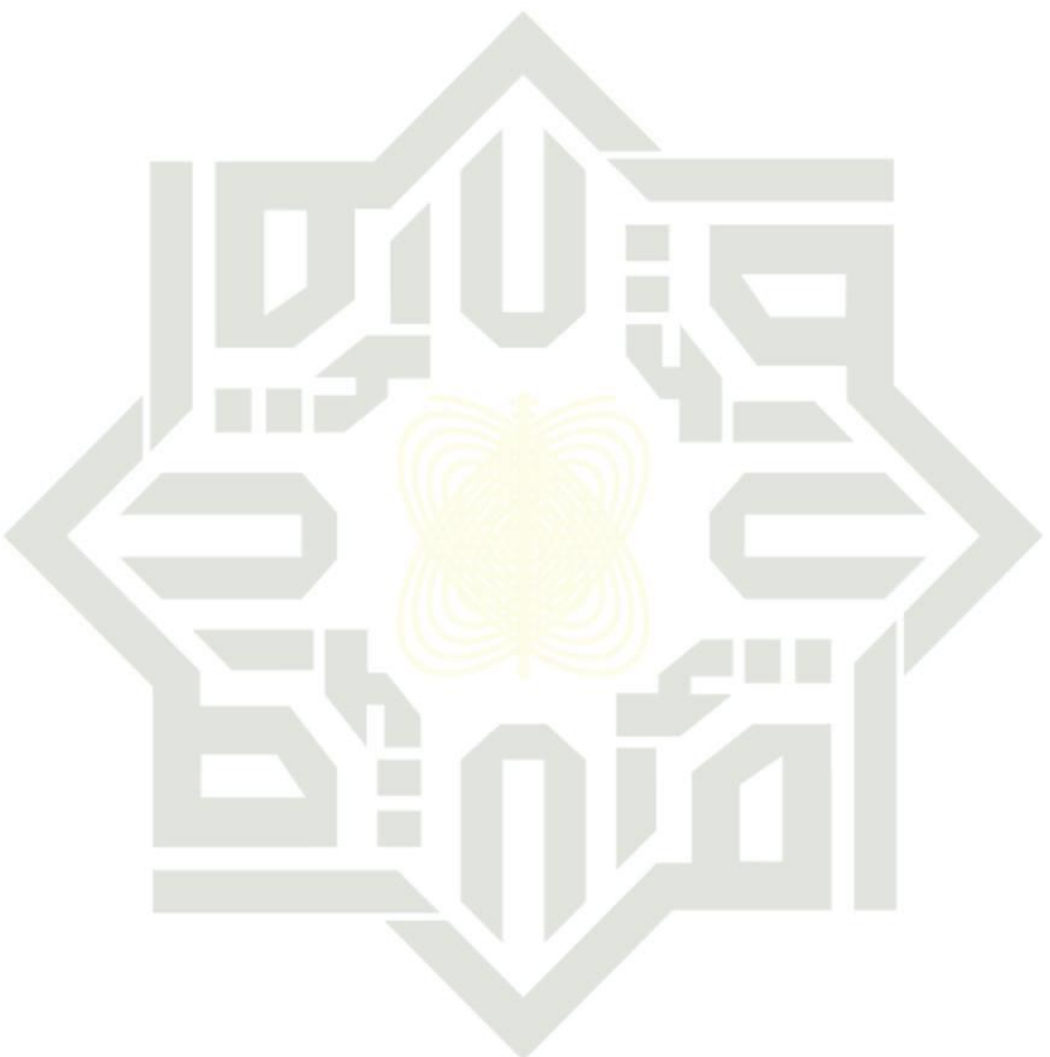
Tahapan pengujian ini dilakukan untuk melihat kualitas serta kekuatan dari sebuah cluster, seberapa baikkah suatu objek tersebut ditempatkan dalam suatu cluster tersebut.

3.8 Kesimpulan

Tahap terakhir yang harus ada pada suatu penelitian adalah kesimpulan dan saran. Pada tahap ini peneliti memberikan kesimpulan dan saran dari penelitian yang telah dilakukannya. Secara umum, kesimpulan menceritakan tentang hasil akhir yang didapatkan pada suatu penelitian, apakah hasil dari penelitian yang dilakukan sesuai dengan apa yang diprediksi di awal atau tidak dan tingkat keberhasilan pada suatu penelitian yang dilakukan juga dapat diketahui pada kesimpulan. Berbeda dengan kesimpulan, saran diberikan oleh

State Islamic University of Sultan Syarif Kasim Riau

peneliti kepada peneliti lainnya tentang perbaikan penelitian yang diharapkan dapat menghasilkan rekomendasi yang lebih baik lagi.



UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB VI PENUTUP

6.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan dapat diambil kesimpulan sebagai berikut.

1. Penerapan algoritma K-Means untuk meng-cluster data Penentuan kelas BPJS Kesehatan didapatkan hasil 3 cluster dengan pengujian menggunakan metode pengujian *Elbow* dan didapatkan hasil 2 cluster menggunakan metode *Silhouette Coefficient*.
2. Berdasarkan hasil dari pengujian dengan menggunakan metode *Elbow* maka dapat disimpulkan bahwa peserta BPJS dikelas 1 mayoritas memiliki rumah kontrak dengan gaji dibawah 3.000.000 dan memiliki jumlah anggota keluarga lebih dari 3. Peserta BPJS dikelas 2 mayoritas memiliki rumah sendiri dengan gaji rata-rata 2.000.000. Sedangkan Peserta BPJS dikelas 3 memiliki rumah sendiri dan memiliki gaji rata-rata diatas 3.000.000.
3. Berdasarkan hasil pengujian menggunakan metode *Elbow* dengan mencari nilai SSE (*Sum of Square of Error*) didapatkan hasil 25 data pada *cluster* 3 dengan selisih nilai SSE 5,98; 35 data pada *cluster* 3 dengan selisih nilai SSE 11,80; 45 data pada *cluster* 3 dengan selisih nilai SSE 16,01; 50 data pada *cluster* 3 dengan selisih nilai SSE 16,83; 60 data pada *cluster* 3 dengan selisih nilai SSE 14,79.
4. Metode *silhouette coefficient* hasil perhitungan menggunakan metode *silhouette coefficient* dari tabel dengan uji coba data sebanyak 20 data dan 5 cluster menghasilkan cluster terbaiknya terdapat di 2 cluster dengan nilai *Si* mendekati angka 1 dan terdapat salah satu hasil *Si* sama dengan 1, dengan nilai *Si* 0,625; 0,25; 1

6.2 Saran

Untuk penelitian selanjutnya saya menyarankan untuk memperhatikan pada proses transformasi dan pengujian.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR PUSTAKA

- Algoritma, P., Data, K., Medis, R., & Diseases, C. (2017). *JURNAL RESTI*. 1(2), 82–89.
- Anggara, M., Sujiani, H., & Helfi, N. (2016). Pemilihan Distance Measure Pada K-Means Clustering Untuk Pengelompokkan Member Di Alvaro Fitness. *Jurnal Sistem Dan Teknologi Informasi*, 1(1), 1–6.
- Bastian, A., Sujadi, H., & Febrianto, G. (n.d.). *Penerapan Algoritma K-Means Clustering Analysis Pada Penyakit Menular Manusia (Studi Kasus Kabupaten Majalengka)*. 1, 26–32.
- Beasiswa, S. P., Musthafa, A., Suyono, H., & Sarosa, M. (2015). *Perbandingan Kinerja Algoritma C . 45 dan AHP-*. 9(2), 109–114.
- Hartanti, N. T. (2020). Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 6(2), 82–89. <https://doi.org/10.25077/teknosi.v6i2.2020.82-89>
- Issn, I. P. E.-. (2018). *Computer Based Information System Journal PENERAPAN PENERIMA UANG DAN BERAS Rika Harman*. 01, 48–55.
- Messerli, F. H., & Devereux, R. B. (1983). Introduction: Left ventricular hypertrophy-Good or evil? *The American Journal of Medicine*, 75(3 PART A), 1–3. [https://doi.org/10.1016/0002-9343\(83\)90110-9](https://doi.org/10.1016/0002-9343(83)90110-9)
- Nasari, F., & Sianturi, C. J. M. (2016). Penerapan Algoritma K-Means Clustering Untuk Pengelompokkan Penyebaran Diare Di Kabupaten Langkat. *CogITO Smart Journal*, 2(2), 108. <https://doi.org/10.31154/cogito.v2i2.19.108-119>
- Slitonga, P. D. P., & Damanik, R. (2016). Implementasi Algoritma K-Means Clustering Pada Analisis Penyebaran Penyakit Pasien Pengguna Badan Penyelenggara Jaminan Sosial Kesehatan (BPJS). *Jurnal Teknik Informatika UNIKA Santo Thomas*, 1(2), 37–41.
- Sudrajat, J., & Pudjiantoro, T. H. (2008). Aplikasi Penterjemah Bahasa Indonesia

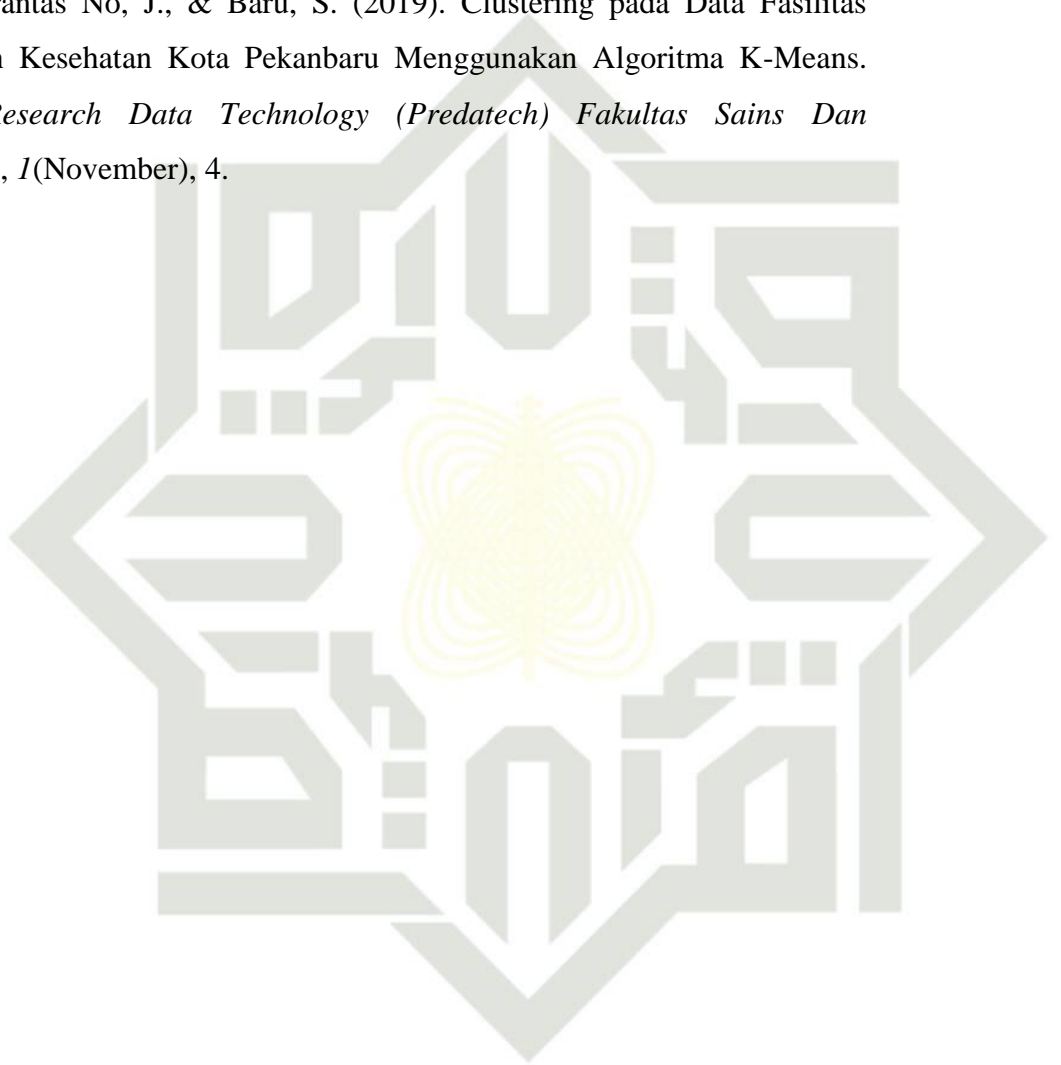


Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Ke Bahasa Inggris Dengan Menggunakan Metode Terselia (Supervised Learning). *Jurnal Computech & Bisnis*, 2(1), 1–13.

Syahidatul Helma, S., Rustiyan, R. R., Normala, E., Studi Sistem Informasi Fakultas Sains dan Teknologi, P., Islam Negeri Sultan Syarif Kasim Riau, U., Soebrantas No, J., & Baru, S. (2019). Clustering pada Data Fasilitas Pelayanan Kesehatan Kota Pekanbaru Menggunakan Algoritma K-Means. *Puzzle Research Data Technology (Predattech) Fakultas Sains Dan Teknologi*, 1(November), 4.



UIN SUSKA RIAU